

CLAIMS

We claim:

1. In a computerized environment, a method of normalizing document data to improve the results of search requests, the method comprising the acts of:

receiving a document containing document data;

parsing the document data into one or more document segments;

identifying at least one of the one or more document segments as an alias that

correlates with a document datum found in an alias directory service; and

associating the received document with the document alias so that, upon request

for the document datum through a search engine, the received document

is returned to the requester by association of the document datum with

the alias.

2. The method of claim 1, wherein the document data are metadata, and wherein the alias is a document metadatum.

3. The method of claim 1, further comprising identifying a secondary document reference contained within the received document; parsing the secondary document into secondary document segments; identifying the secondary document segment with a secondary alias, and associating the secondary document segment with the secondary alias, the secondary document, and the received document.

4. The method of claim 1, wherein the alias directory service is a contact database containing one or more aliases for one or more terms associated with one or more corresponding contacts.

5. The method of claim 1, further comprising identifying the document segment as part of a predefined class or a class alias, so that a data request through a search engine returns the requested data to the requester when the requester enters one or more of the identified class, the class alias, and the alias.

6. The method of claim 5, wherein the class is one or more of a weighted value for one or more associated terms, a metadata concept, and a property type.

7. The method of claim 5, further comprising associating a term in an inverted index with one or more of the identified alias and the predefined class or class alias; and storing the inverted index for use by a search engine.

8. The method of claim 6, wherein the property type is an authorship property.

9. The method of claim 6, wherein a classification module further implements the method comprising:

identifying a next document containing next document data that can be identified with the class, whereby the class comprises at least the document containing document data and the next document containing next document data; and

based on the document data and the next document data, identifying additional documents within the class, so that the classification module is trained to associate additional documents with the class that would not have otherwise been identified.

10. In a computerized environment, a method of normalizing document data to improve the results of search requests, the method comprising:

an act of receiving a document containing document data;

an act of parsing the document data into one or more document segments; and

a step for normalizing document metadata used as a reference by a search engine

by maintaining one or more relationships between a search term and an alternate search term, a search term property or alternative search term property.

11. The method of claim 10, wherein the step for improving future search results returned to a requester of a requested term includes:

an act of identifying at least one of the document segments as an alias for a document datum found in an alias directory service; and

an act of associating the received document with the document alias so that, upon request for the document datum through a search engine, the received document is returned to the requester by association with the alias.

12. The method of claim 10, further comprising receiving, by a gatherer module, directory service data that include one or more aliases for a metadatum.

13. The method of claim 12, further comprising parsing the directory service data; and associating the parsed data with one or more classes so that one or more documents are related by one or more corresponding classes and one or more metadata aliases.

14. The method of claim 13, wherein the directory service data are contained in one or more of a contact database and a text file having delimited values, wherein the delimited values equate one or more alternative terms for a normalized value.

WORKMAN, NYDEGGER & SEELEY
A PROFESSIONAL CORPORATION
ATTORNEYS AT LAW
1000 EAGLE GATE TOWER
60 EAST SOUTH TEMPLE
SALT LAKE CITY, UTAH 84111

15. A computer program product having computer-executable instructions for performing a method of normalizing document data to improve the results of search requests, the method comprising the acts of:

receiving a document containing document data;

parsing the document data into one or more document segments;

identifying at least one of the one or more document segments as an alias for a document datum found in an alias directory service; and

associating the received document with the document alias so that, upon request for the document datum through a search engine, the received document is returned to the requester by association of the document datum with the alias.

16. The computer program product of claim 15, wherein the document data are metadata, and wherein the alias is a document metadatum.

17. The computer program product of claim 15, further comprising identifying a secondary document reference contained within the received document; parsing the secondary document into secondary document segments; identifying the secondary document segment with a secondary alias, and associating the secondary document segment with the secondary alias, the secondary document, and the received document.

18. The computer program product of claim 15, wherein the alias directory service is a contact database containing one or more aliases for one or more terms associated with one or more corresponding contacts.

19. The computer program product of claim 15, further comprising identifying the document segment as part of a predefined class or a class alias, so that a data request through a search engine returns the requested data to the requester when the requester enters one or more of the identified class, the class alias, and the alias.

20. The computer program product of claim 19, wherein the class is one or more of a weighted value for one or more associated terms, a metadata concept, and a property type.

21. The computer program product of claim 19, further comprising associating a term in an inverted index with one or more of the identified alias and the predefined class or class alias; and storing the inverted index for use by a search engine.

22. The computer program product of claim 20, wherein the property type is an authorship property.

23. The computer program product of claim 20, wherein a classification module further implements the method comprising:

identifying a next document containing next document data that can be identified with the class, whereby the class comprises at least the document containing document data and the next document containing next document data; and

based on the document data and the next document data, identifying additional documents within the class, so that the classification module is trained to associate additional documents with the class that would not have otherwise been identified.

24. A computer program product having computer-executable instructions for performing a method of normalizing document data to improve the results of search requests, the method comprising:

an act of receiving a document containing document data;

an act of parsing the document data into one or more document segments; and

a step for normalizing document metadata used as a reference by a search engine by maintaining one or more relationships between a search term and an alternate search term, a search term property or alternative search term property.

25. The computer program product of claim 24, wherein the step for improving future search results returned to a requester of a requested term includes:

an act of identifying at least one of the document segments as an alias for a document datum found in an alias directory service; and

an act of associating the received document with the document alias so that, upon request for the document datum through a search engine, the received document is returned to the requester by association with the alias.

26. The computer program product of claim 24, further comprising receiving, by a gatherer module, directory service data that include one or more aliases for a metadatum.

27. The computer program product of claim 26, further comprising parsing the directory service data; and associating the parsed data with one or more classes so that

one or more documents are related by one or more corresponding classes and one or more metadata aliases.

28. The computer program product of claim 27, wherein the directory service data are contained in one or more of a contact database and a text file having delimited values, wherein the delimited values equate one or more alternative terms for a normalized value.

WORKMAN, NYDEGGER & SEELEY
A PROFESSIONAL CORPORATION
ATTORNEYS AT LAW
1000 EAGLE GATE TOWER
60 EAST SOUTH TEMPLE
SALT LAKE CITY, UTAH 84111